

Pedestrian Detection for UAVs Using Cascade Classifiers and Saliency Maps

Wilbert G. Aguilar^{1,3,4(✉)}, Marco A. Luna², Julio F. Moya²,
Vanessa Abad⁵, Hugo Ruiz^{1,6}, Humberto Parra^{1,7}, and Cecilio Angulo⁴

¹ Dep. Seguridad y Defensa,
Universidad de las Fuerzas Armadas ESPE, Sangolquí, Ecuador
wgaguilar@espe.edu.ec

² Dep. Eléctrica y Electrónica,
Universidad de las Fuerzas Armadas ESPE, Sangolquí, Ecuador

³ CICTE Research Center,
Universidad de las Fuerzas Armadas ESPE, Sangolquí, Ecuador

⁴ GREC Research Group, Universitat Politècnica de Catalunya,
Barcelona, Spain

⁵ Universitat de Barcelona, Barcelona, Spain

⁶ PLM Research Center, Purdue University, West Lafayette, IN, USA

⁷ Universidad Politécnica de Madrid, Madrid, Spain

Abstract. In this paper, we proposed algorithm and dataset for pedestrian detection focused on applications with micro multi rotors UAV (Unmanned Aerial Vehicles). For training dataset we capture images from surveillance cameras at different angles and altitudes. We propose a method based on HAAR-LBP (Local Binary Patterns) cascade classifiers with Adaboost (Adaptive Boosting) training and, additionally we combine cascade classifiers with saliency maps for improving the performance of the pedestrian detector. We evaluate our dataset by the implementation of the HOG (Histogram of oriented gradients) algorithm with Adaboost training and, finally, algorithm performance is compared with other approaches from the state of art. The results shows that our dataset is better for pedestrian detection in UAVs, HAAR-LBP have better characteristics than HAAR like features and the use of saliency maps improves the performance of detectors due to the elimination of false positives in the image.

Keywords: HAAR · HOG · LBP · Saliency maps · People detection · Cascade classifiers · UAVs

1 Introduction

In the field of computer vision, there are several applications in object detection, one of these applications is pedestrian detection, used on surveillance [1, 2], robotics [3, 4, 5, 6], navigation [7, 8, 9], driver assistance systems particularly in PPSs (pedestrian protection systems) [10, 11], and others. In the state of art multiple feature extraction algorithms working with machine learning and datasets have been created to deal with this problem.

Developments in computer vision have been introduced for UAVs [12, 13, 14]. Pedestrian detection can be used with UAVs taking into consideration that they have a complex dynamic and altitude variation that adding extra challenges to the detection [15, 16]. Conventional classifiers fail when altitude increases generating more false positives.

Our proposal for pedestrian detection in UAVs considers the altitude and introduces the CICTE-PeopleDetection dataset with images captured from surveillance cameras. We use two trained algorithms: The first one based on a combination of the feature extraction methods HAAR-LBP, and the second one based on HOG. Both algorithms use cascade classifiers with Adaboost training. In addition we propose an algorithm that merges Saliency Maps algorithm presented on [17] with cascade classifier to provide detection robustness. Our proposal is evaluated in images captured from UAVs in different scenarios.

This paper is organized as follows: Sect. 2 describes the related work on pedestrian detection. Next, our proposal for pedestrian detection, the creation of dataset and the algorithm are described in Sect. 3. In Sect. 4 we present the experimental results, followed by the summary. Finally conclusions and future works are presented in Sect. 5.

2 Related Works

In the literature, several research groups have created different datasets and methods for pedestrian detection. INRIA was introduced on [18], with training based on Histograms of Oriented Gradients (HOG). Widely used datasets are Caltech Pedestrian Dataset [19] and KITTI [20], due to they are comparatively large and challenging. According to [21, 22] there are two types of datasets: photo datasets and video datasets. Photo datasets like MIT [23], CVC [11], NICTA [24] abord the classification problem: train binary classification algorithms. Video datasets as ETH [25], TUD-Brussels [26] or Dalmier (DB) [27] are focused on detection problem: design and test full image detection systems and human locomotion modeling.

Two important algorithms have been developed for pedestrian detection and object detection in general: Haar-like features [28] by Viola and Jones, and Dalal and Triggs algorithm called HOG [18]. Both algorithms have generated over 40 new approaches [21]. Several methods for pedestrian detection includes feature extraction algorithms: HAAR [28], HOG [18, 29], HOG-HAAR [30] and HOG-LBP [31]; working with machine learning approaches based on SVMs [18, 32] or Adaboost [11, 27].

The applications of pedestrian detection in UAVs are manifold: Human safety [33], rescue and monitoring missions [34, 35], track people systems [32, 36], and others. One of the challenges of pedestrian detection in UAVs is the camera perspective variations that deform the images. In [37, 38], they use thermal imagery combined with cascade classifiers to perform the detection. Few papers like [35] works on altitudes around five meters. In this paper, authors propose post-disaster victims detection with cascade classifier methods. In UAVs, the use of saliency maps is widely used to object and motion detection in aerial images [35, 39]. Works like [34] use saliency maps to detect people reducing the search space, choosing randomly bounding boxes to detect

people inside saliency region and treating separately all detection windows; they fuse the results using mean-shift procedure applied in flights from 10 to 40 m of altitude.

3 Our Approach

3.1 Dataset Creation

One of the reasons for introducing our dataset is the requirement to detect people from UAV cameras. The main problem in pedestrian detection is the high altitude, where people images have deformation of their characteristics. The main difference of CICTE-PeopleDetection with previous photo datasets is the location and perspective of the cameras that emulate the onboard camera perspective of the UAV. We use surveillance cameras for photo dataset creation due to UAVs video captures are stable and comparable with fixed cameras. There are approximately 100 cameras (we can not specify the exactly number of cameras for security reasons) with D1 resolution located in the University between 2.3 m and 5 m of height looking down as shown in Fig. 1.



Fig. 1. Location of the cameras in the campus (a) 5 m height. (b) 2.3 m

For training we need positive and negative images. Positive images are the images that contain the object to be detected, in our case pedestrians. Negative images are frames without pedestrians. Our dataset has 3900 positive images and 1212 negative images. The positives images were captured in the Universidad de las Fuerzas Armadas ESPE during the day and the night in different scenarios, and contain entire and partial occluded people samples.

3.2 Training Process

Our approach consists in the combination of two algorithms for extraction of the feature set: Local Binary Patterns (LBP) and Haar-like features. We use Adaptive Boosting (AdaBoost) as training algorithm and a combination of Haar-LBP features due to them are algorithms of low computation time. To create our Haar-LBP algorithm we divided the all images in 70% for training and the other 30% for testing, after that we use the algorithm with a UAV images in different scenarios. Additionally, we train a HOG cascade classifier and compare it with Opencv HOG to validate our Dataset. The training processes are shown Fig. 2.

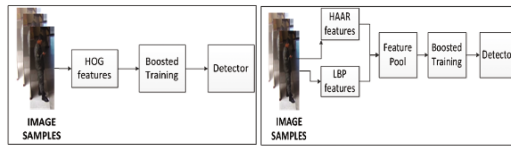


Fig. 2. Pedestrian detection training. (a) HOG features with Adaboost. (b) Haar-LBP features with Adaboost

The methods used for training the cascade classifiers are described as follows:

Local Binary Patterns (LBP)

This feature extractor was presented in [40] as a texture descriptor for object detection, and compares a central pixel with the neighbours. The window to be examined is separated into cells of 16×16 pixels. 8 neighbours are considered for each pixel inside the cell, the central pixel value is the threshold. A value of 1 is assigned if the neighbour is greater or equal to the central pixel, otherwise the value is 0.

Haar-like Features

Viola and Jones uses a statistical approach for the tracking and detection problem, describing the ratio between light and dark areas within a defined kernel. This algorithm is robust regarding to noise and lighting changes. The method uses simple feature sets similar to Haar basis functions [28, 41].

Histogram of Oriented Gradients (HOG)

This algorithm is a feature descriptor for object detection focused on pedestrian detection and introduced in [18]. The image window is separated into smaller parts called cells. For each cell, we accumulate a local 1-D histogram of gradient orientations of the pixels in the cell. Each cell is discretized into angular bins according to the gradient orientation and each pixel of the cell contributes with a gradient weight to its corresponding angular bin. The adjacent cells are grouped in special regions called blocks and the normalized group of histograms represents the block histogram.

Adaboost

Adaboost is a machine learning algorithm [42] that initially keeps uniform distribution of weights in each training sample. In the first iteration the algorithm trains a weak classifier using a feature extraction methods or mix of them achieving a higher recognition performance for the training samples. In the second iteration, the training samples, misclassified by the first weak classifier, receive higher weights. The new selected feature extraction methods should be focused in these misclassified samples.

3.3 People Detection Algorithm

In order to get a better performance of the classifier we implement a combination of cascade classifier with saliency maps, an algorithm presented in [17]. The purpose of saliency maps is to locate prominent areas at every location in the visual field. The areas with high saliency correspond to objects or places they are most likely to be found, and the areas with lower saliency are associated to background [43]. The saliency maps

algorithms are deduced by convolving the function f by an isotropic bi-dimensional Gaussian function [44]:

$$S(X) = f(X)G_{\sigma}(X) \quad (1)$$

where σ is the standard deviation of the Gaussian function. The standard deviation depends on the experimental setup (size of the screen and viewing distance). To eliminate the false positives in the image we obtain the salient region; we consider a threshold from the salient map and we create a mask where values greater than threshold will belong to salient map. Additionally, this region was dilated to give it robustness. This algorithm is shown in Fig. 3.

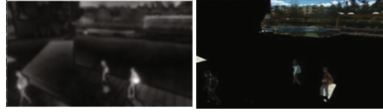


Fig. 3. Saliency maps algorithm. (a) Saliency map (b) Saliency region.

Once it has been obtained the salient region, our algorithm proposes take as true positives only the cascade classifier detections inside this region. For this reason we take the salient region as Region of Interest (ROI). To determinate if a detection bounding box is inside the salient region, we compute the center point of the bounding box with the formulas:

$$x_m = x + \left(\frac{w}{2}\right); \quad y_m = y + \left(\frac{h}{2}\right) \quad (2)$$

where x and y are the horizontal and vertical coordinates of the top left of the bounding box, x_m and y_m are the coordinates of the central point and w, h are the width and height. We take the center point as reference to avoid false positives that could have small parts of their bounding box in salient regions. Unlike other methods presented in the literature [34], we use our own algorithm for combination of cascade classifier with saliency maps. Our proposal is presented graphically in Fig. 4.

The results of the application of this algorithm are presented in the Sect. 4.

4 Results and Discussion

4.1 Dataset and Training Evaluation

The metric of evaluation for our approach is based on the sensitivity (true positive rate-TPR) and the miss rate (False negative rate-FNR). Defined as follows:

$$TPR = \frac{TP}{TP + FN} * 100\%; \quad FNR = \frac{FN}{TP + FN} * 100\% \quad (3)$$

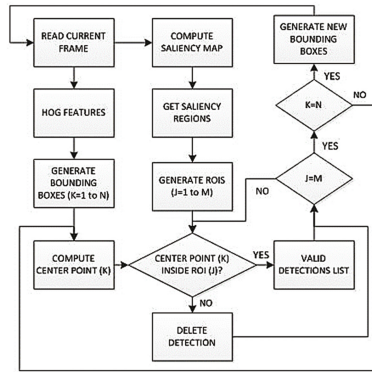


Fig. 4. Algorithm for people detection using HOG cascade classifier and saliency Maps

For the dataset evaluation we have trained the cascade classifier based on HOG features and compared this classifier with the OpenCV HOG cascade classifier. We tested the cascade classifier with videos captured from UAVs. Experimental results are presented in Table 1.

Table 1. Dataset training performance

Algorithm	Sensitivity (%)	Miss rate (%)
HOG-CICTE	55.71 %	44.28%
HOG [45]	35.71%	64.28%

In this table, two cascade classifiers are compared: HOG-CICTE PeopleDetection and a HOG cascade classifier with the Adaboost training from the OpenCV library. Result shows our approach has better performance, the miss rate of our proposal is 20% lower than the conventional classifier miss rate, and the sensitivity is higher. ROC curves for comparing both algorithms are presented in Fig. 5.

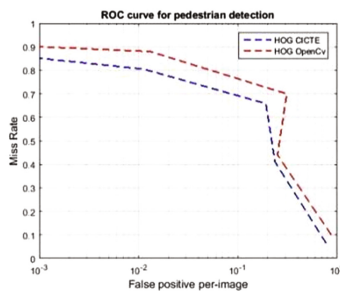


Fig. 5. Comparison of ROC curves for HOG from OpenCV cascade classifier and HOG-CICTE cascade classifier trained for pedestrian detection in UAVs.

In Fig. 5, HOG-CICTE classifier has a better performance that HOG from OpenCV cascade classifier in videos captured from UAVs.

4.2 Algorithm Evaluation

For the algorithm evaluation we are using 3 scenarios with 3 different altitudes. We compare HAAR-LBP features and HOG features (trained with CICTE-PeopleDetection) respect to other cascade classifiers. Results are presented in Table 2.

Table 2. Cascade classifiers performance

Algorithm	Altitude (m)	Sensitivity (%)	Miss rate (%)
HAAR-LBP	2	10.66%	89.34%
	3	7.18%	92.82%
	4	5.43%	94.57%
HAAR [28]	2	3.27%	96.73%
	3	2.28%	97.72%
	4	1.68%	98.32%
LBP [46]	2	41.98%	58.02%
	3	33.18%	66.82%
	4	23.09%	76.91%
HOG CICTE	2	67.23%	32.76%
	3	63.72%	36.27%
	4	60.44%	39.56%

In Table 2, the combination of HAAR-LBP features has low sensitivity compared with the other methods; however the proposal is higher than HAAR features. With altitude increasing, sensitivity decrease in all cascade classifiers. Performance curves are presented in the Fig. 6.

In the Fig. 6, the performance of the HAAR-LBP features algorithm is better than HAAR individually applied. HAAR-LBP features generate a lower rate of false positives. True positive rate of HAAR-LBP features is higher than HAAR features but

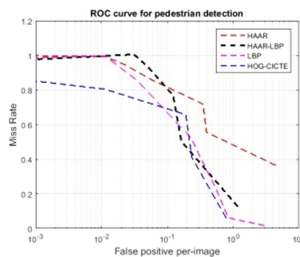


Fig. 6. Comparison of ROC curves for different approaches.

lower than LBP. Nevertheless, the HOG-CICTE cascade classifier still has the best performance due to its higher true positives rate and lower false positives rate.

4.3 Cascade Classifier-Saliency Maps Combination

Based on the results of performance we choose HOG CICTE cascade classifier to implement our algorithm. Graphical results are shown in the Fig. 7 and video results are provided by: https://www.youtube.com/watch?v=KN_hVgp1_t4

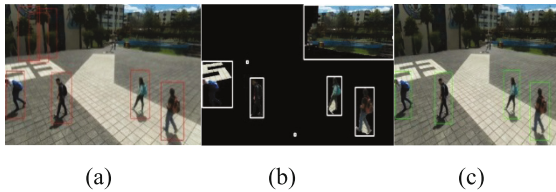


Fig. 7. Combination of Saliency Maps and Cascade classifier. (a) Cascade classifier result (b) saliency regions (c) final Result

As we can see in the Fig. 7, the use of saliency region helps to reject the false positives in the images. For the evaluation we take an additional metric of evaluation that is precision or positive predictive value (PPV), given by:

$$PPV = \frac{TP}{TP + FP} * 100\%$$

where TP are the true positive values and FP are the false positives. The results of precision of the detector with the application of the saliency region algorithm (SR) are shown in the Table 3.

Table 3. Comparison of precision between algorithms

Algorithm	Precision (%)
HOG-CICTE	72.23%
HOG-CICTE +SR	92.1%

In Table 3, the application of saliency region algorithm improves the precision of detection in 20% approximately, this denote an improvement in the performance too. The performance curves of two algorithms are shown in the Fig. 8.

The Fig. 8 shows that the use of saliency region algorithm improves the detection performance eliminating false positives.

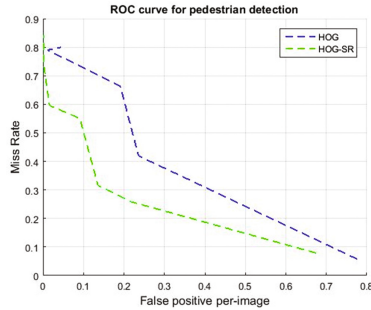


Fig. 8. ROC curves for HOG and HOG-SR

5 Conclusions and Future Work

Our proposal for pedestrian detection based on HOG features has higher performance than OpenCV HOG respect to sensitivity and miss rate (with an improvement of 20%), as shown in the Table 1 and Fig. 5, because the images used for training emulate UAVs perspective.

In order to improve the HAAR algorithm performance we combine two algorithms (HAAR and LBP). The sensitivity increased and the miss rate decreased as shows Table 2 and Fig. 6; however the performance is lower in comparison with HOG-CICTE and LBP algorithm. When the altitude increased from 2 to 4 meters, the sensitivity decreased in the four algorithms. Comparing HAAR-LBP and HAAR, HAAR-LBP has a better performance even in the altitude of 4 m.

The use of saliency maps improves the performance detectors, saliency map helps to eliminate background regions even in mobile cameras like UAVs, and these regions may contain objects that confuse the classifier that is important to decrease the number of false positives.

In the future is necessary to improve the detection. We will train new classifiers with images captured from UAVs, taking into consideration other human body parts like face, head, shoulders, etc. In addition, a robust of detection could be used for many applications like people tracking or people avoidance systems.

Acknowledgement. This work is part of the projects VisualNavDrone 2016-PIC-024 and MultiNavCar 2016-PIC-025, from the Universidad de las Fuerzas Armadas ESPE, directed by Dr. Wilbert G. Aguilar.

References

1. Torresan, H.: Advanced surveillance systems: combining video and thermal imagery for pedestrian detection. In: Proceedings of SPIE, pp. 506–515 (2004)
2. Zhang, L.Z.L., Wu, B.W.B., Nevatia, R.: Pedestrian detection in infrared images based on local shape features. In: 2007 IEEE Conference on Computer Vision and Pattern Recognition, pp. 1–8 (2007)

3. Aguilar, W.G., Angulo, C., Costa, R., Molina, L.: Control autónomo de cuadricopteros para seguimiento de trayectorias. In: *Memorias del IX Congreso de Ciencia y Tecnología ESPE 2014* (2014)
4. Aguilar, W.G., Angulo, C.: Compensación de los Efectos Generados en la Imagen por el Control de Navegación del Robot Aibo ERS 7. In: *Memorias del VII Congreso de Ciencia y Tecnología, ESPE 2012*, pp. 165–170, June 2012
5. Jafari, O.H., Mitzel, D., Leibe, B.: Real-time RGB-D based people detection and tracking for mobile robots and head-worn cameras. In: *Proceedings of IEEE International Conference on Robotics and Automation*, pp. 5636–5643, April 2016
6. Kobilarov, M., Sukhatme, G., Hyams, J., Batavia, P.: People tracking and following with mobile robot using an omnidirectional camera and a laser. In: *2006 IEEE International Conference on Robotics and Automation 2006 ICRA*, pp. 557–562, May 2006
7. Aguilar, W.G., Casaliglla, V., Pólit, J.: Obstacle avoidance based-visual navigation for micro aerial vehicles. *Electronics* **6**(1), 10 (2017)
8. Cabras, P., Rosell, J., Pérez, A., Aguilar, W.G., Rosell, A.: Haptic-based navigation for the virtual bronchoscopy. In: *18th IFAC World Congress, Milano, Italy* (2011)
9. Aguilar, W.G., Morales, S.: 3D environment mapping using the Kinect V2 and path planning based on RRT algorithms. *Electronics* **5**(4), 70 (2016)
10. Gavrilu, D.M.: Pedestrian detection from a moving vehicle. In: *Proceedings of the 6th European Conference on Computer Vision*, vol. 1843, pp. 37–49 (2000)
11. Gerónimo, D., López, A.M., Sappa, A.D., Graf, T.: Survey of pedestrian detection for advanced driver assistance systems. *IEEE Trans. Pattern Anal. Mach. Intell.* **32**(7), 1239–1258 (2010)
12. Aguilar, W.G., Angulo, C.: Real-time video stabilization without phantom movements for micro aerial vehicles. *EURASIP J. Image Video Process.* **1**, 1–13 (2014)
13. Aguilar, W.G., Angulo, C.: Real-time model-based video stabilization for microaerial vehicles. *Neural Process. Lett.* **43**(2), 459–477 (2016)
14. Aguilar, W.G., Angulo, C.: Robust video stabilization based on motion intention for low-cost micro aerial vehicles. In: *2014 11th International Multi-Conference on Systems, Signals Devices (SSD)*, pp. 1–6 (2014)
15. Rudol, P., Doherty, P.: Human Body Detection and Geolocalization for UAV Search and Rescue Missions Using Color and Thermal Imagery
16. Aguilar, W.G., Luna, M.A., Moya, J.F., Abad, V., Parra, H., Ruiz, H.: Pedestrian detection for UAVs using cascade classifiers with meanshift. In: *2017 IEEE 11th International Conference on Semantic Computing (ICSC)*, pp. 509–514 (2017)
17. Itti, L., Koch, C., Niebur, E.: A model of saliency-based visual attention for rapid scene analysis. *IEEE Trans. Pattern Anal.* **20**, 1254–1259 (1998)
18. Dalal, N., Triggs, W.: Histograms of oriented gradients for human detection. In: *2005 IEEE Computer Society Conference on Computer Vision Pattern Recognition CVPR 2005*, vol. 1, no. 3, pp. 886–893 (2004)
19. Dollár, P., Wojek, C., Schiele, B., Perona, P.: Pedestrian detection: a benchmark. In: *2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, CVPR Workshops 2009*, pp. 304–311 (2009)
20. Geiger, A., Lenz, P., Urtasun, R.: Are we ready for autonomous driving? The KITTI vision benchmark suite. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 3354–3361 (2012)
21. Benenson, R., Omran, M., Hosang, J., Schiele, B.: Ten years of pedestrian detection, what have we learned? In: *Proceedings of Computer Vision-ECCV 2014 Workshop*, pp. 613–627 (2014)

22. Dollár, P., Wojek, C., Schiele, B., Perona, P.: Pedestrian detection: an evaluation of the state of the art. *IEEE Trans. Pattern Anal. Mach. Intell.* **34**(4), 743–761 (2012)
23. Papageorgiou, C., Poggio, T.: Trainable system for object detection. *Int. J. Comput. Vis.* **38**(1), 15–33 (2000)
24. Overett, G., Petersson, L., Brewer, N., Andersson, L., Pettersson, N.: A new pedestrian dataset for supervised learning. In: *Proceedings of IEEE Intelligent Vehicles Symposium*, pp. 373–378 (2008)
25. Ess, A., Leibe, B., Schindler, K., van Gool, L.: Robust multiperson tracking from a mobile platform. *IEEE Trans. Pattern Anal. Mach. Intell.* **31**(10), 1831–1846 (2009)
26. Wojek, C., Walk, S., Schiele, B.: Multi-Cue onboard pedestrian detection. In: *2009 IEEE Computer Society Conference on Computer Vision Pattern Recognition Workshops CVPR Workshops 2009*, pp. 794–801 (2009)
27. Enzweiler, M., Gavrila, D.M.: Monocular pedestrian detection: survey and experiments. *IEEE Trans. Pattern Anal. Mach. Intell.* **31**(12), 2179–2195 (2009)
28. Viola, P., Jones, M.: Rapid object detection using a boosted cascade of simple features. In: *Conference on Computer Vision Pattern Recognition*, pp. 1–9 (2001)
29. Zhu, Q., Avidan, S., Yeh, M.C., Cheng, K.T.: Fast human detection using a cascade of histograms of oriented gradients. *IEEE Conf. Comput. Vis. Pattern Recognit.* **2**, 1491–1498 (2006)
30. Wojek, C., Schiele, B.: A performance evaluation of single and multi-feature people detection. In: *Pattern Recognition Symposium*, pp. 82–91 (2008)
31. Wang, X., Han, T.X., Yan, S.: An HOG-LBP human detector with partial occlusion handling. In: *2009 IEEE 12th International Conference on Computer Vision ICCV*, pp. 32–39 (2009)
32. Imamura, Y., Okamoto, S., Lee, J.H.: Human tracking by a multi-rotor drone using HOG features and linear SVM on images captured by a monocular camera. In: *Proceedings of the International MultiConference of Engineers and Computer Scientists*, vol. 1, pp. 8–13 (2016)
33. Lioulemes, A., Galatas, G., Metsis, V., Mariottini, G.L., Makedon, F.: Safety challenges in using AR. drone to collaborate with humans in indoor environments. In: *Proceedings of 7th International Conference on Pervasive Technologies Related to Assistive Environments*, p. 33 (2014)
34. Blondel, P., Potelle, A., Pegard, C., Lozano, R.: Human detection in uncluttered environments: from ground to UAV view. In: *2014 13th International Conference on Control Automation Robotics and Vision ICARCV 2014*, pp. 76–81 (1997)
35. Andriluka, M., Schnitzspan, P., Meyer, J., Kohlbrecher, S., Petersen, K., Von Stryk, O., Roth, S., Schiele, B.: Vision based victim detection from unmanned aerial vehicles. In: *2010 IEEE/RSJ International Conference on Intelligent Robot and System (IROS)*, pp. 1740–1747, October 2010
36. De Smedt, F., Hulens, D., Goedeme, T.: On-board real-time tracking of pedestrians on a UAV. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshop*, pp. 1–8, October 2015
37. Rudol, P., Doherty, P., Science, I.: Human body detection and geolocalization for UAV search and rescue missions using color and thermal imagery. In: *2008 IEEE Aerospace Conference*, pp. 1–8 (2008)
38. Gąszczak, A., Breckon, T.P., Han, J.: Real-time people and vehicle detection from UAV imagery. In: *IS&T/SPIE Electron. Imaging*, pp. 8–11, January 2011
39. Siam, M., Elhelw, M.: Robust autonomous visual detection and tracking of moving targets in UAV imagery. In: *Proceedings of International Conference on Signal Process (ICSP)*, vol. 2, pp. 1060–1066, December (2012)

40. Wang, L., He, D.: Texture classification using texture spectrum. *Pattern Recognit.* **23**, 905–910 (1990)
41. Papageorgiou, C.P. Oren, M.: A general framework for object detection. In: *IEEE International Conference on Computer Vision*, pp. 555–562, January 1998
42. Schapire, R.E., Singer, Y.: Improved boosting algorithms using confidence-rated predictions. *Mach. Learn.* **37**(3), 297–336 (1999)
43. Moosmann, F., Larlus, D., Jurie, F.: Learning saliency maps for object categorization. In: *International Workshop on the Representation and Use of Prior Knowledge in Vision* (2006)
44. Le Meur, O., Baccino, T.: Methods for comparing scanpaths and saliency maps: strengths and weaknesses. *Behav. Res. Methods* **45**, 251–266 (2013)
45. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: *2005 IEEE Computer Society Conference on Computer Vision Pattern Recognition*, vol. 1, pp. 886–893 (2005)
46. Ojala, T., Pietikäinen, M., Mäenpää, T.: A generalized local binary pattern operator for multiresolution gray scale and rotation invariant texture classification. *Adv. Pattern Recognit.* **2013**, 399–408 (2001)